ENTROPY OF CONVEX FUNCTIONS ON $\mathbb{R}^d$

FUCHANG GAO AND JON A. WELLNER

ABSTRACT. Let Ω be a bounded closed convex set in $\mathbb{R}^d$ with non-empty interior, and let $\mathcal{C}_r(\Omega)$ be the class of convex functions on Ω with L^r-norm bounded by 1. We obtain sharp estimates of the ε-entropy of $\mathcal{C}_r(\Omega)$ under $L^p(\Omega)$ metrics, $1 \le p < r \le \infty$. In particular, the results imply that the universal lower bound $\varepsilon^{-d/2}$ is attained by all d-polytopes. While a general convex body can be approximated by inscribed polytopes, the entropy rate does not carry over to the limiting body. For example, if Ω is the closed unit ball in $\mathbb{R}^d$, then the metric entropy of $\mathcal{C}_\infty(\Omega)$ under the $L^p(\Omega)$ metrics has order $\varepsilon^{-(d-1)p/2}$ for $p > \frac{d}{d-1}$. Our results have applications to questions concerning rates of convergence of nonparametric estimators of high-dimensional shape-constrained functions.

1. Introduction

Convex functions are of special importance not only because they are basic classes of functions, but also because they appear so commonly in applications. For instance, in nonparametric estimation of densities in statistics many interesting classes of densities are defined in terms of convex or concave functions on a bounded convex region.

Let Ω be a compact convex set in $\mathbb{R}^d$ with non-empty interior. For $1 < r \le \infty$, let

$$\mathcal{C}_r(\Omega) = \left\{ f \;\middle|\; f \text{ is convex on } \Omega, \left(\int_\Omega |f(x)|^r dx\right)^{1/r} \le 1 \right\}.$$

In this paper, we are interested in the covering number $N(\varepsilon, \mathcal{C}_r(\Omega), \|\cdot\|_p)$ of $\mathcal{C}_r(\Omega)$ under $L^p(\Omega)$ distances, where $1 \le p < r$, which is defined as the minimum number of closed balls of radius ε in $L^p(\Omega)$ distance that are needed to cover $\mathcal{C}_r(\Omega)$. When $r = \infty$, for the purpose of statistical applications, we will also consider the bracketing entropy number $N_{[\,]}(\varepsilon, \mathcal{C}_\infty(\Omega), \|\cdot\|_p)$ of $\mathcal{C}_\infty(\Omega)$ under $L^p(\Omega)$ distance, which is defined as the minimum number of ε-brackets

$$[\underline{f}, \overline{f}] := \left\{ g \in \mathcal{C}_\infty(\Omega) \;\middle|\; \underline{f} \le g \le \overline{f} \right\}, \; \|\overline{f} - \underline{f}\|_p \le \varepsilon$$

needed to cover $\mathcal{C}_\infty(\Omega)$. It is easy to see that

$$N(\varepsilon, \mathcal{C}_\infty(\Omega), \|\cdot\|_{L^p(\Omega)}) \le N_{[\,]}(2\varepsilon, \mathcal{C}_\infty(\Omega), \|\cdot\|_{L^p(\Omega)}).$$

2010 *Mathematics Subject Classification.* Primary 52A41, 41A46; Secondary 52A27, 52C17, 52B11.

Key words and phrases. metric entropy, bracketing entropy, convex functions, polytopes, simplicial approximation.

Research partially supported by a grant from the Simons Foundation, #246211.

Supported in part by NSF Grant DMS-1104832 and NI-AID grant 2R01 AI291968-04.

In the one dimensional case, a sharp estimate of $\log N(\varepsilon, \mathcal{C}_\infty([0,1]), \|\cdot\|_2)$ was obtained in [5] by connecting with the small ball probability of integrated Brownian motions. In fact, it was shown in [5] that for all bounded k monotone function classes $\mathcal{M}_k([0,1])$ for all $k \geq 1$, we have

$$\log N(\varepsilon, \mathcal{M}_k([0,1]), \|\cdot\|_{L^2([0,1])}) \asymp \varepsilon^{-1/k},$$

where

$$\mathcal{M}_k([0,1]) = \left\{ f \;\middle|\; -1 \leq f(x) \leq 1, (-1)^i f^{(i)}(x) \geq 0, 1 \leq i \leq k, x \in [0,1] \right\}.$$

Note that since $\mathcal{C}_\infty([0,1]) \subset \mathcal{M}_2([0,1]) - \mathcal{M}_2([0,1])$, the result above immediately implies that

$$\log N(\varepsilon, \mathcal{C}_\infty([0,1]), \|\cdot\|_2) \asymp \varepsilon^{-1/2}.$$

A constructive proof with sharp estimate of $N_{[\,]}(\varepsilon, \mathcal{M}_k([0,1]), \|\cdot\|_p)$ for all k-monotone functions, $k \geq 1$ was given in [6]. The estimate of $\log N(\varepsilon, \mathcal{C}_\infty([0,1]), \|\cdot\|_2)$ was later rediscovered by [3].

In high dimensional case, if $\Omega = [0,1]^d$, Guntuboyina and Sen proved that

$$\log N(\varepsilon, \mathcal{C}_r([0,1]^d), \|\cdot\|_p) \asymp \varepsilon^{-d/2}$$

for both $r = \infty$ [9] and $r > p$ [8]. Besides the hypercubes, there seem to be no other results available.

While the set $[0,1]^d$ maintains the simple geometric structure similar to that of $[0,1]$, other shapes are much harder to deal with because their lack of nice geometric structure. Let us point out that for $d \geq 2$, even if Ω is convex, the rate of metric entropy of $\mathcal{C}_r(\Omega)$ could heavily depend on the shape of the convex body Ω. For example, our Theorem 1.5 below says that when Ω is the closed unit ball of $\mathbb{R}^d$, the entropy of $\mathcal{C}_\infty(\Omega)$ has a growth rate far larger than $\varepsilon^{-d/2}$.

In this paper, we study classes of convex functions on all compact convex sets Ω in $\mathbb{R}^d$ with non-empty interior. We first show that $\varepsilon^{-d/2}$ is the general lower bound, and if Ω is a convex polytope, then $\varepsilon^{-d/2}$ is also the upper bound. More precisely, we will prove the following theorem.

Theorem 1.1. *Let Ω be a compact convex set in $\mathbb{R}^d$ with non-empty interior. Let $\mathcal{C}_r(\Omega)$ be the set of convex functions on Ω whose $L^r(\Omega)$-norms are bounded by 1.*

(i) There exists a constant c_1 is a constant depending only on d such that for all $\varepsilon > 0$,

$$\log N(\varepsilon, \mathcal{C}_r(\Omega), \|\cdot\|_{L^p(\Omega)}) \geq c_1 |\Omega|^{\frac{d}{2p} - \frac{d}{2r}} \varepsilon^{-d/2}.$$

(ii) If Ω can be triangulated into m simplices of dimension d, then any $1 \leq p < r$, there exists a constant C_1 depending on p, d, r, such that for any $\varepsilon > 0$,

$$\log N(\varepsilon, \mathcal{C}_r(\Omega), \|\cdot\|_{L^p(\Omega)}) \leq C_1 m |\Omega|^{\frac{d}{2p} - \frac{d}{2r}} \varepsilon^{-d/2}.$$

Consequently, if Ω is a convex polytope with v extreme points, then we can choose $m = O(v^{\lceil \frac{d}{2} \rceil})$. When $r = \infty$, the same inequality holds for bracketing entropy.

In view of the fact that a general compact convex set can be approximated by convex sets with finitely many extreme points, one might guess that the rate $\varepsilon^{-d/2}$ holds for general compact convex sets in $\mathbb{R}^d$ with non-empty interior. This, however, is not the case. This is because the upper bound increases linearly as m increases. This linear dependence on m is important. It enables us to establish upper bounds for general bounded convex sets. For that, we need the following definition.

Definition 1.2. Let Ω be a bounded closed set in $\mathbb{R}^d$ with non-empty interior. A sequence of $d-$simplices $\{D_1, D_2, \ldots\}$ is called an *admissible simplicial approximation sequence* for Ω if $D_i \subset \Omega$ for all $i \in \mathbb{N}$ and $D_i^\circ \cap D_j^\circ = \emptyset$ for all $i \neq j$ (where D° denotes the interior of D). Let $\mathcal{D}_\Omega$ denote the collection of all admissible simplicial approximation sequences for Ω. For an admissible approximation sequence $\{D_i\} \in \mathcal{D}_\Omega$ and $t \in (0,1)$ we define

$$S(t, \Omega; \{D_i\}) \equiv \min\{j \in \mathbb{N} : \ |\Omega \setminus \cup_{i \le j} D_i| \le t|\Omega|\},$$

and we call

$$S(t, \Omega) \equiv \inf\{S(t, \Omega; \{D_i\}) : \ \{D_i\} \ \in \mathcal{D}_\Omega\}$$

the *simplicial approximation number* of Ω.

Now we can state the following theorem.

Theorem 1.3. *Let Ω be a compact convex set in $\mathbb{R}^d$ with non-empty interior. Let $\mathcal{C}_r(\Omega)$ be the set of convex functions on Ω whose $L^r(\Omega)$-norms are bounded by 1. Then, there exists a constant C depending only on d, p and r, such that for any $0 < \varepsilon < 1$,*

$$\begin{aligned} \log N(\varepsilon|\Omega|^{\frac{1}{p}-\frac{1}{r}}, \mathcal{C}_r(\Omega), \|\cdot\|_{L^p(\Omega)}) \quad &\le \quad C \int_{\delta(\varepsilon)}^1 \frac{S(t,\Omega)}{t} dt \\ &\quad + C \left(\int_{\delta(\varepsilon)}^1 \left(\frac{S(t,\Omega)}{t} \right)^\beta dt \right)^{1/\beta} \cdot \varepsilon^{-d/2} \end{aligned}$$

where $\delta(\varepsilon) = 2^{-2-\frac{r}{r-p}} \varepsilon^{\frac{rp}{r-p}}$ and $\beta = \frac{2pr}{2pr+(r-p)d}$.

As an example, we consider the case when Ω is the closed unit ball in $\mathbb{R}^d$. By specifically constructing a simplicial sphere approximation for the ball, we show that Theorem 1.3 implies the following corollary. Our proof of the corollary also provides a general scheme of constructing simplicial sphere approximations for uniformly smooth convex bodies.

Corollary 1.4. *If Ω is the closed unit ball in $\mathbb{R}^d$, then there exists a constant C depending only on p and d such that for all $0 < \varepsilon < 1$,*

$$\log N(\varepsilon, \mathcal{C}_\infty(\Omega), \|\cdot\|_{L^p(\Omega)}) \le \begin{cases} C\varepsilon^{-\frac{(d-1)p}{2}} & \text{if } p > \frac{d}{d-1} \\ C\varepsilon^{-d/2} |\log \varepsilon|^{\frac{d+1}{2}} & \text{if } p = \frac{d}{d-1} \\ C\varepsilon^{-d/2} & \text{if } p < \frac{d}{d-1} \end{cases}.$$

The following theorem implies the sharpness of Theorem 1.3: at least for the case when Ω is the closed unit ball in $\mathbb{R}^d$, and $p \neq \frac{d}{d-1}$, the upper bound in Corollary 1.4 is optimal, which in turn implies that the upper bound in Theorem 1.3 is optimal, and the dependence on m in Theorem 1.1 (ii) is optimal.

Theorem 1.5. *If Ω is the closed unit ball in $\mathbb{R}^d$, then there exists a constant c_2 dependent only on d and p such that for all $0 < \varepsilon < 1$,*

$$\log N(\varepsilon, \mathcal{C}_\infty(\Omega), \|\cdot\|_{L^p(\Omega)}) \ge c_2 \varepsilon^{-\beta},$$

where $\beta = \max\{(d-1)p/2, d/2\}$.

2. Proofs

2.1. **Scaling.** In this subsection, we prove two lemmas, through which we can reduce a problem on an arbitrary closed convex set with non-empty interior to a problem on a closed convex set contained in $[0,1]^d$ with volume at least $1/d!$.

Lemma 2.1 (Boxing a Convex Set). *Every compact convex set Ω in $\mathbb{R}^d$ with a non-empty interior can be enclosed in a closed rectangular box of volume $d!|\Omega|$, and contains a convex polytope of at most $2d$ vertices and volume at least $|\Omega|/d!$, where $|\Omega|$ stands for the Lebesgue measure of Ω.*

Proof. We use induction on d to show that we can find positive numbers $h_1, h_2, \ldots, h_d$ such that Ω is contained in a rectangular box of size $h_1 \times h_2 \times \cdots \times h_d$, and contains a convex polytope of at most $2d$ vertices with volume at least $\frac{1}{d!} \cdot h_1 \times h_2 \times \cdots \times h_d$.

The statement is trivial if $d = 1$. Suppose the statement is true for $d = k$. Consider the case $d = k + 1$. Let $h_{k+1} = \mathrm{diam}(\Omega)$. Choose $x, y \in \Omega$ so that $\|x - y\| = h_{k+1}$. Let $P_x^{\perp}(\Omega)$ be the projection of Ω onto the affine hyperplane that contains x and is orthogonal to $x - y$. Since $P_x^{\perp}(\Omega) \subset \mathbb{R}^k$ is a k-dimensional compact convex set with non-empty interior, by the induction hypothesis, we can find positive numbers $h_1, h_2, \ldots, h_k$ such that $P_x^{\perp}(\Omega)$ is contained in a rectangular box R_k of size $h_1 \times h_2 \times \cdots \times h_d$, and contains a convex polytope T_k of at most $2k$ vertices with volume at least $\frac{1}{k!} \cdot h_1 \times h_2 \times \cdots \times h_k$. If we let $[x, y]$ be the line segment between x and y, then Ω is clearly contained in the rectangular box $R_k \times [x, y]$ of size $h_1 \times h_2 \times \cdots \times h_{k+1}$.

To show that Ω contains a convex polytope of at most $2(k + 1)$ vertices with volume at least $\frac{1}{(k+1)!} \cdot h_1 \times h_2 \times \cdots \times h_{k+1}$, we let $u_1, u_2, \ldots u_m$, $m \le 2k$, be the vertices of the convex polytope T_k. Clearly, the convex hull U of $\{x, y, u_1, u_2, \ldots, u_m\}$ has volume $|U| \ge \frac{1}{(k+1)!} \cdot h_1 \times h_2 \times \cdots \times h_{k+1}$.

For each $1 \le i \le m$, there exists $z_i \in \Omega$ such that $P_x^{\perp} z_i = u_i$. Because Ω is convex, it contains the convex hull of $\{x, y, z_1, z_2, \cdots, z_m\}$. Denote this convex hull by T_{k+1}. Then T_{k+1} has at most $2(k + 1)$ vertices. Note that the volume of T_{k+1} is at least as large as $|U|$. Indeed, for any unit vector u perpendicular to $x - y$, consider the half-line in the direction of u starting from x. Suppose the half-line intersects the boundary U at $w(u)$. Choose $z(u) \in T_{k+1}$ such that $P_x^{\perp} z(u) = w(u)$. Clearly, the area of $\triangle xyz(u)$ is the same as that of $\triangle xyw(u)$, which equals $\frac{1}{2} h_{h+1} \|x - w(u)\|_2$. Let σ_{k-1} be the $(k - 1)$-dimensional spherical measure. By using a cylindrical system to compute the volume of T_{k+1}, we have

$$|T_{k+1}| \ge \int_{S^{k-1}} \mathrm{area}(\triangle xyz(u))\, d\sigma_{k-1}(u) = \int_{S^{k-1}} \mathrm{area}(\triangle xyw(u))\, d\sigma_{k-1}(u) = |U|.$$

Hence, $|T_{k+1}| \ge \frac{1}{(k+1)!} \cdot h_1 \times h_2 \times \cdots \times h_{k+1}$. This proves the case $d = k + 1$, and thus the lemma. □

Lemma 2.2 (Scaling). *Let Ω be a bounded closed convex set contained in a closed rectangular box R with volume $|R|$, and let T be any affine transform that maps R onto $[0,1]^d$. Then for all $1 \le p < r < \infty$ and $\varepsilon > 0$,*

$$N(\varepsilon, \mathcal{C}_r(\Omega), \|\cdot\|_{L^p(\Omega)}) = N(|R|^{\frac{1}{r} - \frac{1}{p}} \varepsilon, \mathcal{C}_r(T(\Omega)), \|\cdot\|_{L^p(T(\Omega))}). \tag{2.1}$$

Similarly, for all $1 \le p < \infty$ and $\varepsilon > 0$,

$$N_{[\,]}(\varepsilon, \mathcal{C}_\infty(\Omega), \|\cdot\|_{L^p(\Omega)}) = N_{[\,]}(|R|^{-\frac{1}{p}} \varepsilon, \mathcal{C}_\infty(T(\Omega)), \|\cdot\|_{L^p(T(\Omega))}). \tag{2.2}$$

Proof. Let $f \in \mathcal{C}_r(T(\Omega))$. Then $|R|^{-1/r} f \circ T \in \mathcal{C}_r(\Omega)$ since

$$\int_\Omega |f|^r \circ T d\lambda = \int_{T(\Omega)} |f|^r |T^{-1}| d\lambda = |R| \int_{T(\Omega)} |f|^r d\lambda \le |R|.$$

Now let $f_1, \ldots, f_N$ be an $L_p(\Omega)$ ε-net for $\mathcal{C}_r(\Omega)$. Then, for $f \in \mathcal{C}_r(T(\Omega))$ we have

$$\begin{aligned} &\left(\int_{T(\Omega)} \left| f - |R|^{1/r} f_i \circ T^{-1} \right|^p d\lambda \right)^{1/p} \\ &= \left(\int_\Omega \left| f \circ T - |R|^{1/r} f_i \right|^p |T| d\lambda \right)^{1/p} \\ &= \left(\int_\Omega \left| |R|^{1/r} \left(|R|^{-1/r} f \circ T - f_i \right) \right|^p |R|^{-1} d\lambda \right)^{1/p} \\ &= |R|^{1/r - 1/p} \left(\int_\Omega \left| |R|^{-1/r} f \circ T - f_i \right|^p d\lambda \right)^{1/p}, \end{aligned}$$

and since $|R|^{-1/r} f \circ T \in \mathcal{C}_r(\Omega)$, for some $i \in \{1, \ldots, N\}$ the last display is bounded above by $|R|^{1/r-1/p}\varepsilon$. Thus given an $L_p(\Omega)$ ε-net for $\mathcal{C}_r(\Omega)$ we have constructed an $L_p(T(\Omega))$ $|R|^{1/p-1/r}\varepsilon$-net for $\mathcal{C}_r(T(\Omega))$. It follows that

$$N(|R|^{1/r-1/p}\varepsilon, \mathcal{C}_r(T(\Omega)), \|\cdot\|_{L_p(T(\Omega))}) \le N(\varepsilon, \mathcal{C}_r(\Omega), \|\cdot\|_{L_p(\Omega)}). \tag{2.3}$$

By a similar argument we find that

$$N(|R|^{1/r-1/p}\varepsilon, \mathcal{C}_r(T(\Omega)), \|\cdot\|_{L_p(T(\Omega))}) \ge N(\varepsilon, \mathcal{C}_r(\Omega), \|\cdot\|_{L_p(\Omega)}), \tag{2.4}$$

and hence the equality (2.1) holds.

To prove (2.2), first note that if $f \in \mathcal{C}_\infty(T(\Omega))$, then $\sup_\Omega |f \circ T| = \sup_{T(\Omega)} |f| \le 1$, so $f \circ T \in \mathcal{C}_\infty(\Omega)$. Then suppose that $[\underline{f}_i, \overline{f}_i]$, $1 \le i \le N$, are $L_p(\Omega)$ brackets of size ε for $\mathcal{C}_\infty(\Omega)$. Then $[\underline{f}_i \circ T^{-1}, \overline{f}_i \circ T^{-1}]$, $1 \le i \le N$, are $L_p(T(\Omega))$ brackets of size $|R|^{-1/p}\varepsilon$ for $\mathcal{C}_\infty(T(\Omega))$. To see this, note that for some $i \in \{1, \ldots, N\}$

$$\underline{f}_i(x) \le f \circ T(x) \le \overline{f}_i(x) \quad \text{for all} \quad x \in \Omega,$$

and hence

$$\underline{f}_i \circ T^{-1}(y) \le f(y) \le \overline{f}_i \circ T^{-1}(y) \quad \text{for all} \quad y \in T(\Omega).$$

Furthermore,

$$\begin{aligned} \int_{T(\Omega)} |\overline{f}_i \circ T^{-1}(y) - \underline{f}_i \circ T^{-1}(y)|^p d\lambda &= \int_\Omega |\overline{f}_i - \underline{f}_i|^p |T| d\lambda \\ &= |R|^{-1} \int_\Omega |\overline{f}_i - \underline{f}_i|^p d\lambda \le \left(|R|^{-1/p}\varepsilon \right)^p. \end{aligned}$$

Thus

$$N_{[\,]}(|R|^{-1/p}\varepsilon, \mathcal{C}_\infty(T(\Omega)), \|\cdot\|_{L_p(T(\Omega)}) \le N_{[\,]}(\varepsilon, \mathcal{C}_\infty(\Omega), \|\cdot\|_{L_p(\Omega)}).$$

A similar argument yields the reversed inequality, and hence (2.2) holds. □

By combining Lemma 2.1 and Lemma 2.2, we have

$$N(\varepsilon, \mathcal{C}_r(\Omega), \|\cdot\|_{L^p(\Omega)}) \le N((d!)^{\frac{1}{r}-\frac{1}{p}} \cdot |\Omega|^{\frac{1}{r}-\frac{1}{p}}\varepsilon, \mathcal{C}_r(T(\Omega)), \|\cdot\|_{L^p(T(\Omega))}), \tag{2.5}$$

and

$$N_{[\,]}(\varepsilon, \mathcal{C}_\infty(\Omega), \|\cdot\|_{L^p(\Omega)}) \le N_{[\,]}((d!)^{-\frac{1}{p}} \cdot |\Omega|^{-\frac{1}{p}}\varepsilon, \mathcal{C}_\infty(T(\Omega)), \|\cdot\|_{L^p(T(\Omega))}), \tag{2.6}$$

where $T(\Omega)$ has volume at least $1/d!$ and is contained in $[0,1]^d$.

2.2. **Under Uniform Lipschitz.** In this subsection, we recall that if we assume the functions in $\mathcal{C}_r(\Omega)$ are bounded and uniform Lipschitz, then the metric entropy estimate would follow from the following known results of Bronshtein [1].

Lemma 2.3 (Bronshtein)**.** *Let $\mathcal{K}(\rho)$ be the set of all closed convex sets contained in the closed Euclidean ball of radius ρ in $\mathbb{R}^{d+1}$, $d \geq 1$. Let h be the Hausdorff distance on $\mathcal{K}(\rho)$. There exists a constant C_0 depending only on d, such that for any $0 < \varepsilon < \rho$,*

$$\log N(\varepsilon, \mathcal{K}(\rho), h) \leq C_0(\rho\varepsilon^{-1})^{d/2}.$$

Lemma 2.4. *Let Ω be a closed convex set in $[0,1]^d$, and let $\mathcal{F}_\alpha(\Omega)$ be the class of convex functions on Ω that are bounded by M and have Lipschitz constant bounded by α. Then for all $\varepsilon < 2^{-1-1/p}\sqrt{(1+\alpha^2)(M^2+d/4)}$,*

$$\log N_{[\,]}(\varepsilon, \mathcal{F}_\alpha(\Omega), \|\cdot\|_{L^p(\Omega)}) \leq 2^{-d} C_0\{(1+\alpha^2)(4M^2+d)\}^{d/4}\varepsilon^{-d/2}$$

where C_0 is the same constant as in Lemma 2.3.

Remark 2.5. Lemma 2.3 can be found in [1], [4], or [14] Lemma 2.7.8, page 163. Lemma 2.4 is also known. For example, it would follow from [14] Corollary 2.7.10, page 164. Because we deal with bracketing entropy, we include a proof here for the convenience of the reader.

Proof. For each $f \in \mathcal{F}_\alpha$, since Ω is a closed and convex set and f is convex, the epigraph $\mathrm{epi}(f) \equiv \{(x,t): \ f(x) \leq t \leq M, \ x \in \Omega\}$ is a closed convex set contained in the closed Euclidean ball in $\mathbb{R}^{d+1}$ with radius $\sqrt{d/4 + M^2}$ and center at $(1/2, 1/2, \ldots, 1/2, 0)$.

On the other hand, for any $x \in \Omega$, $y \in \Omega$, and $f, g \in \mathcal{F}_\alpha$,

$$\begin{aligned}
|f(x) - g(x)| &\leq |f(x) - f(y)| + |f(y) - g(x)| \\
&\leq \alpha\|x - y\|_2 + |f(y) - g(x)| \\
&\leq \sqrt{1+\alpha^2}\|(x, g(x)) - (y, f(y))\|_2.
\end{aligned}$$

Taking the infimum on $y \in \Omega$ followed by the supremum on $x \in \Omega$, we find that

$$\|f - g\|_\infty \leq \sqrt{1+\alpha^2}\, h(\mathrm{epi}(f), \mathrm{epi}(g)).$$

Thus, by Lemma 2.3

$$\begin{aligned}
\log N(\eta, \mathcal{C}_\infty(\Omega), \|\cdot\|_{L^\infty(\Omega)}) &\leq \log N((1+\alpha^2)^{-1/2}\eta, \mathcal{K}(\sqrt{M^2 + d/4}), h) \\
&\leq C_0\{\sqrt{(1+\alpha^2)(M^2+d/4)}\eta^{-1}\}^{d/2}.
\end{aligned}$$

Thus there exist $N \leq \exp(C_0\{\sqrt{(1+\alpha^2)(M^2+d/4)}\eta^{-1}\}^{d/2})$ functions $f_1, \ldots, f_N$ defined on Ω, such that for each $f \in \mathcal{K}(\Omega)$, there exists some f_i, $i \in \{1, 2, \ldots, N\}$, such that $|f(x) - f_i(x)| \leq \eta$ for all $x \in \Omega$. For each $i \in \{1, \ldots, N\}$ define

$$\begin{aligned}
\overline{f}_i(x) &= \sup\{f(x) : \ |f(x) - f_i(x)| \leq \eta, \ f \in \mathcal{F}_\alpha\}; \\
\underline{f}_i(x) &= \inf\{f(x) : \ |f(x) - f_i(x)| \leq \eta, \ f \in \mathcal{F}_\alpha\}
\end{aligned}$$

for each $x \in \Omega$. Then we have

$$\|\overline{f}_i - \underline{f}_i\|_{L^\infty(\Omega)} \leq 2\eta.$$

In particular this implies that for all $1 \le p < \infty$

$$\int_\Omega |\overline{f}_i(x) - \underline{f}_i(x)|^p dx \le (2\eta)^p.$$

Letting $\varepsilon = 2\eta$ we find that

$$\begin{aligned}\log N_{[\,]}(\varepsilon, \mathcal{C}_\infty(\Omega), \|\cdot\|_{L^p(\Omega)}) &\le \log N(2\eta, \mathcal{C}_\infty(\Omega), \|\cdot\|_\infty)\\ &\le C_0\{\sqrt{(1+\alpha^2)(M^2+d/4)}(2\eta)^{-1}\}^{d/2}\\ &= 2^{-d} C_0\{(1+\alpha^2)(4M^2+d)\}^{d/4}\varepsilon^{-d/2}.\end{aligned}$$

□

2.3. **Paring the Boundary.** In this subsection, we show that if we pare off the boundary of Ω by δ, and consider the set

$$\Omega_\delta = \{x \in \Omega \mid \operatorname{dist}(x, \partial\Omega) \ge \delta\}, \tag{2.7}$$

then considering functions restricted to Ω_δ, the entropy can be estimated using Lemma 2.4. The details are proved in the following two lemmas.

Lemma 2.6. *Let Ω be a compact convex set in $[0,1]^d$ with $|\Omega| \ge 1/d!$. Then there exists a constant Λ depending only on d, such that for any $1 \le r \le \infty$ and any $0 < \delta \le 1$,*

$$\mathcal{C}_r(\Omega) \subset \Lambda\delta^{-d/r} \cdot \mathcal{C}_\infty(\Omega_\delta),$$

where Ω_δ is as defined in (2.7). In fact $\Lambda = \max\{(d\Gamma(d/2)/\pi^{d/2})^{1/r}, (d!)d2^{d+2}\}$ works.

Proof. First, we show that if $f \in \mathcal{C}_r(\Omega)$, then on Ω

$$f \ge -(d!)^{1/r} 2^{d+2} d. \tag{2.8}$$

Let x_0 be a minimizer of f on Ω. If $f(x_0) \ge 0$, then there is nothing to prove; otherwise, the set $K := \{x \in \Omega \mid f(x) \le 0\}$ is a closed convex set with x_0 as an interior point. Denote $K_0 = K - x_0$, and define

$$K_\eta = \Omega \cap \{x_0 + [(1+\eta)K_0 \setminus (1-\eta)K_0]\},$$

where $0 < \eta < 1$. We show that if $x \notin K_\eta$, then $|f(x)| > \eta|f(x_0)|$. Indeed, consider a function g on Ω defined so that: $g(x_0) = f(x_0)$, $g(\gamma) = f(\gamma)$ for all $\gamma \in \partial K$, and g is linear on line segment

$$L_\gamma := \{x \in \Omega \mid x = x_0 + t(\gamma - x_0), t \ge 0\}.$$

Then, by the convexity of f on each L_γ, we have $|f(x)| \ge |g(x)|$ on Ω. Because for all $x \notin K_\eta$, $\|x - \gamma\| \ge \eta\|x_0 - \gamma\|$, we have

$$|g(x)| = |g(\gamma)| + \frac{\|x-\gamma\|}{\|x_0 - \gamma\|}|f(x_0)| > \eta|f(x_0)|.$$

Hence, on $\Omega \setminus K_\eta$, $|f(x)| \ge \eta|f(x_0)|$.

Because the volume of K_η is bounded by $[(1+\eta)^d - (1-\eta)^d] \cdot |K| \le d2^d\eta|\Omega|$, we have

$$1 \ge \int_{\Omega\setminus K_\eta} |f(x)|^r dx \ge (\eta|f(x_0)|)^r \cdot [1 - d2^d\eta] \cdot |\Omega|.$$

This implies that

$$|f(x_0)| \le \eta^{-1}|\Omega|^{-1/r}(1 - d2^d\eta)]^{-1/r}.$$

By choosing $\eta = [d2^d(1+1/r)]^{-1}$, we obtain

$$|f(x_0)| \le |\Omega|^{-1/r} d2^d \left(1+\frac{1}{r}\right)(1+r)^{1/r} \le (d!)^{1/r} 2^{d+2} d \le (d!) 2^{d+2} d.$$

This proves (2.8).

Next, we show that there exists a constant C depending on Ω such that on Ω_δ, $f(x) \le \Lambda \delta^{-d/r}$.

Let z_0 be a maximizer of f on Ω_δ. If $f(z_0) \le 0$, there is nothing to prove. So, we assume $f(z_0) > 0$. Let $V = \{x \in \Omega \mid f(x) < f(z_0)\}$. Then V is a convex set with z_0 at its boundary. There exists a hyperplane that separates V and z_0. This hyperplane separates Ω into two parts. On the part not containing V, $f \ge f(z_0)$. In particular, $f \ge f(z_0)$ on the half of the ball centered at z_0 with radius δ. Calling this half of the ball W, we have

$$1 \ge \int_W |f(x)|^r dx \ge \frac{\pi^{d/2}}{d\Gamma(d/2)} \delta^d f(z_0)^r,$$

which implies that

$$f(z_0) \le \left(\frac{d\Gamma(d/2)}{\pi^{d/2}}\right)^{1/r} \delta^{-d/r}.$$

Together with (2.8) we obtain that there exists some Λ depending only on d such that for all $x \in \Omega_\delta$, $|f(x)| \le \Lambda \delta^{-d/r}$. □

Lemma 2.7. *Let Ω be a closed convex set in $[0,1]^d$. For any $1 \le r \le \infty$ and any $0 < \delta \le 1$,*

$$N_{[\,]}(\varepsilon, \mathcal{C}_r(\Omega), \|\cdot\|_{L^p(\Omega_\delta)}) \le \exp\Big(C_2 \delta^{-d/2 - d^2/r} \varepsilon^{-d/2}\Big),$$

where Ω_δ is as defined in (2.7) and C_2 is a constant depending only on d, C_0 in Lemma 2.3, and Λ in Lemma 2.6.

Proof. We show that when restricted to Ω_δ, f has a Lipschitz constant bounded by $2^{2+d/r}\Lambda\delta^{-1-d/r}$. Indeed, by Lemma 2.6, f is bounded by $2^{d/r}\Lambda\delta^{-d/r}$ on $\Omega_{\delta/2}$. Note that $\Omega_\delta \subset \Omega_{\delta/2} \subset \Omega$. Thus by [14], problem 2.7.4 page 165, f is Lipschitz on Ω_δ with Lipschitz constant $2(\delta/2)^{-1}2^{d/r}\Lambda\delta^{-d/r} = 2^{2+d/r}\Lambda\delta^{-1-d/r}$.

Thus by Lemma 2.4 it follows that

$$\begin{aligned}
&\log N_{[\,]}(\epsilon, \mathcal{C}_r(\Omega), \|\cdot\|_{L^p(\Omega_\delta)}) \\
&\qquad \le 2^{-d} C_0 (1 + 2^{4+2d/r}\Lambda^2\delta^{-2-2d/r})^{d/4} (4\Lambda^2\delta^{-2d/r} + d)^{d/4} \epsilon^{-d/2} \\
&\qquad \le C_2 \delta^{-d/2-d^2/r} \varepsilon^{-d/2} \qquad (2.9)
\end{aligned}$$

for some constant C_2 depending only on d and r. □

2.4. **Combining.** In this subsection, we prove some lemmas that enables us to studying metric entropy by decomposing the set Ω.

Lemma 2.8 (Union)**.** *If $\Omega = \cup_{i=1}^k \Omega_i$, then for all $1 \le p < r \le \infty$,*

$$N(\varepsilon, \mathcal{C}_r(\Omega), \|\cdot\|_{L^p(\Omega)}) \le \prod_{i=1}^k N\left(\eta_i, \mathcal{C}_r(\Omega_i), \|\cdot\|_{L^p(\Omega_i)}\right), \qquad (2.10)$$

$$N_{[\,]}(\varepsilon, \mathcal{C}_\infty(\Omega), \|\cdot\|_{L^p(\Omega)}) \le \prod_{i=1}^k N_{[\,]}\left(\eta_i, \mathcal{C}_\infty(\Omega_i), \|\cdot\|_{L^p(\Omega_i)}\right), \qquad (2.11)$$

where $\varepsilon = (\sum_{i=1}^k \eta_i^p)^{1/p}$. *Furthermore, if* $\Omega_1, \Omega_2, \dots, \Omega_k$ *have disjoint interiors, then*

$$N(\varepsilon, \mathcal{C}_r(\Omega), \|\cdot\|_{L^p(\Omega)}) \le 4^k \prod_{i=1}^{k} N\left(\eta_i, \mathcal{C}_r(\Omega_i), \|\cdot\|_{L^p(\Omega_i)}\right) \tag{2.12}$$

where

$$\left(\sum_{i=1}^{k} \eta_i^{\frac{rp}{r-p}}\right)^{\frac{r-p}{rp}} \le 2^{-1/r}\varepsilon.$$

Proof. For each $i \in \{1, 2, \dots, k\}$, there exists a set $\mathcal{N}_i$ of N_i elements, where

$$N_i \equiv N(\eta_i, \mathcal{C}_r(\Omega_i), \|\cdot\|_{L^p(\Omega_i)})$$

such that, for each $f \in \mathcal{C}_r(\Omega) \subset \mathcal{C}_r(\Omega_i)$, there exists $f_i \in \mathcal{N}_i$ satisfying

$$\int_{\Omega_i} |f_i(x) - f(x)|^p dx \le \eta_i^p.$$

Define $\hat{f}(x) = f_i(x)$ for $x \in \Omega_i \setminus \cup_{j<i}\Omega_j$, $1 \le i \le k$. Then we have

$$\int_{\Omega} |f(x) - \hat{f}(x)|^p dx \le \sum_{i=1}^{k} \int_{\Omega_i} |f(x) - f_i(x)|^p dx \le \sum_{i=1}^{k} \eta_i^p = \varepsilon^p.$$

Because there are no more than $N_1 N_2 \cdots N_k$ realizations of $\hat{f}$, (2.10) follows.

The proof of (2.11) is similar. For each $i \in \{1, 2, \dots, k\}$, there exists a set $\mathcal{N}_i$ of $\hat{N}_i$ brackets, where

$$\hat{N}_i \equiv N_{[\,]}(\eta_i, \mathcal{C}_\infty(\Omega_i), \|\cdot\|_{L^p(\Omega_i)})$$

such that, for each $f \in \mathcal{C}_\infty(\Omega) \subset \mathcal{C}_\infty(\Omega_i)$, there exists a bracket $[\underline{f}_i, \overline{f}_i] \in \mathcal{N}_i$ satisfying $\underline{f}_j(x) \le f(x) \le \overline{f}_j(x)$ for all $x \in \Omega_i$, and

$$\int_{\Omega_i} |\overline{f}_i(x) - \underline{f}_i(x)|^p dx \le \eta_i^p.$$

Define $\overline{f}(x) = \overline{f}_i(x)$, $\underline{f}(x) = \underline{f}_i(x)$, $x \in \Omega_i \setminus \cup_{j<i}\Omega_j$, $1 \le i \le k$. Then we have $\underline{f}(x) \le f(x) \le \overline{f}(x)$ for all $x \in \Omega$, and

$$\int_{\Omega} |\overline{f}(x) - \underline{f}(x)|^p dx \le \sum_{i=1}^{k} \int_{\Omega_i} |\overline{f}_i(x) - \underline{f}_i(x)|^p dx \le \sum_{i=1}^{k} \eta_i^p = \varepsilon^p.$$

That is, $[\underline{f}, \overline{f}]$ is an ε-bracket in $L^p(\Omega)$ which contains f. Because there are no more than $\hat{N}_1 \hat{N}_2 \cdots \hat{N}_k$ realizations of $[\underline{f}, \overline{f}]$, (2.11) follows.

Now we turn to the proof of (2.12). For any $f \in \mathcal{C}_r(\Omega)$, and for each $i = 1, 2, \dots, k$, define $n_i(f)$ as the smallest positive integer such that

$$n_i(f) \ge k \int_{\Omega_i} |f(x)|^r dx.$$

Then, $n_i(f) < k \int_{\Omega_i} |f(x)|^r dx + 1$, and using the fact that $\sum_{i=1}^k \int_{\Omega_i} |f|^r dx \le 1$, we get

$$n_1(f) + n_2(f) + \cdots + n_k(f) \le \sum_{i=1}^{k} \left(k \int_{\Omega_i} |f(x)|^r dx + 1\right) \le 2k.$$

Let

$$\mathcal{I} = \left\{(n_1, n_2, \ldots, n_k) \in \mathbb{N}^k \;\middle|\; n_1 + n_2 + \cdots + n_k \le 2k\right\}.$$

For each $I = (i_1, i_2, \ldots, i_k) \in \mathcal{I}$, define

$$\mathcal{F}_I = \{f \in \mathcal{C}_r(\Omega) \mid n_j(f) = i_j, 1 \le j \le k\}.$$

Then we have $\mathcal{C}_r(\Omega) = \cup_{I\in\mathcal{I}}\mathcal{F}_I$. Thus,

$$N(\varepsilon, \mathcal{C}_r(\Omega), \|\cdot\|_{L^p(\Omega)}) \le \sum_{I\in\mathcal{I}} N(\varepsilon, \mathcal{F}_I, \|\cdot\|_{L^p(\Omega)}).$$

Note that for each $j = 1, 2, \ldots, k$, $\mathcal{F}_I \subset (i_j/k)^{1/r}\mathcal{C}_r(\Omega_j)$. Thus,

$$\begin{aligned} N(\eta_j, \mathcal{C}_r(\Omega_j), \|\cdot\|_{L^p(\Omega_j)}) &= N((i_j/k)^{1/r}\eta_j, (i_j/k)^{1/r}\mathcal{C}_r(\Omega_i), \|\cdot\|_{L^p(\Omega_j)}) \\ &\ge N((i_j/k)^{1/r}\eta_j, \mathcal{F}_I, \|\cdot\|_{L^p(\Omega_j)}). \end{aligned}$$

Therefore, for each $1 \le j \le k$, there exists a set $\mathcal{N}_j$ of $Z_j := N(\eta_j, \mathcal{C}_r(\Omega_i), \|\cdot\|_{L^p(\Omega_j)})$ elements such that for each $f \in \mathcal{F}_I$, there exists $f_j \in \mathcal{N}_j$ satisfying

$$\int_{\Omega_j} |f(x) - f_j(x)|^p dx \le (i_j/k)^{p/r}\eta_j^p.$$

If we define $\hat{f}(x) = f_j(x)$ for $x \in \Omega_j \setminus \cup_{r<j}\Omega_r$, then we have

$$\begin{aligned} \int_{\Omega} |f(x) - \hat{f}(x)|^p dx &\le \sum_{j=1}^k (i_j/k)^{p/r}\eta_j^p \\ &\le \left(\sum_{j=1}^k \frac{i_j}{k}\right)^{\frac{p}{r}} \left(\sum_{j=1}^k \eta_j^{\frac{rp}{r-p}}\right)^{1-\frac{p}{r}} \le 2^{\frac{p}{r}} \left(\sum_{j=1}^k \eta_j^{\frac{rp}{r-p}}\right)^{1-\frac{p}{r}} \le \varepsilon^p. \end{aligned}$$

Since, there are no more than $Z_1Z_2\cdots Z_k$ realizations of $\hat{f}$, we obtain

$$N(\varepsilon, \mathcal{F}_I, \|\cdot\|_{L^p(\Omega)}) \le \prod_{i=1}^k N\left(\eta_i, \mathcal{C}_r(\Omega_i), \|\cdot\|_{L^p(\Omega_i)}\right).$$

Note that $\mathcal{I}$ has cardinality $\binom{2k}{k} < 4^k$, and hence (2.12) follows. □

2.5. **With Finitely Many Facets.** Note that by the well-known Upper Bound Theorem of discrete geometry [11], [7], if Ω is a convex polytope with v vertices, then it has at most $k \le 2\binom{v}{\lfloor d/2\rfloor} \le 2v^{\lfloor d/2\rfloor}$ facets. In this subsection, we consider the case when Ω is a convex polytope with k facets. We first prove the upper bound with constant Ck^γ with some $\gamma > 1$. We will use it later only for the case $k = d+1$. However, since the proof is the same, we prove it for the general k.

By scaling, we can assume that Ω is contained in unit d-cube with volume at least $1/d!$. Thus, there exists a point $O \in \Omega$ such that the distance between O and the boundary of Ω is at least $\delta_0 := 1/(2dd!)$. This is because the boundary of $[0, 1]^d$ has $(d-1)$-dimensional area $2d$, and its projection onto Ω is a contraction, thus, the boundary of Ω has $(d-1)$-dimensional area at most $2d$, and by a Bonnesen-style inequality (Corollary 2, page 25 of [12]) the inradius of Ω is at least its volume divided by the $(d-1)$-dimensional surface area of its boundary, i.e. the inradius is at least δ_0.

By otherwise using a translation, we can assume that O is the origin. Let F_i be the i-th facet of Ω for $i = 1, \ldots, k$. Let V_i denote the convex hull of F_i and O. Then,

V_i, $i \in \{1, \ldots, k\}$, form a partition of Ω. For $\delta < \delta_0 \equiv \frac{1}{2d^2 d!}$, let $D_0 := (1-\delta/\delta_0)\Omega$. Define $\Omega_i = V_i \setminus D_0^\circ$, where D_0° denotes the interior of D_0. Then we have

$$\Omega = D_0 \cup \Omega_1 \cup \Omega_2 \cup \cdots \cup \Omega_k.$$

Note that each Ω_i has no more than $k+1$ facets. To see this, we first observe V_i has at most k facets. Indeed, each of the facets of V_i besides F_i is the convex hull of a $(d-2)$-dimensional face of F_i and O. However, each $(d-2)$-dimensional face of F_i corresponds to the intersection of F_i and another facet of Ω. Thus, the number of $(d-2)$-dimensional faces of F_i is at most $k-1$. Therefore, the number of facets of V_i is at most k. Notice that Ω_i has one more facet than V_i. Hence, the number of facets of Ω_i is at most $k+1$. By (2.12) we have

$$N(\varepsilon, \mathcal{C}_r(\Omega), \cdot\|_{L^p(\Omega)}) \le 4^{k+1} N(\eta_0, \mathcal{C}_r(D_0), \|\cdot\|_{L^p(D_0)}) \prod_{i=1}^{k} N(\eta_i, \mathcal{C}_r(\Omega_i), \|\cdot\|_{L^p(\Omega_i)}), \tag{2.13}$$

where $\eta_0 = 2^{-\frac{1}{p}}\varepsilon$, and

$$\eta_i = 2^{-\frac{1}{p}} \left(\frac{|\Omega_i|}{\sum_{i=1}^k |\Omega_i|} \right)^{\frac{1}{p}-\frac{1}{r}} \epsilon.$$

Because $D_0 \subset \Omega_\delta$, by Lemma 2.7, we have

$$\log N(\varepsilon, \mathcal{C}_r(\Omega), \|\cdot\|_{L^p(D_0)}) \le C_2 \delta^{-\frac{d}{2}-\frac{d^2}{r}} (\varepsilon)^{-d/2}.$$

On the other hand, if we let T_i be an affine transform that maps Ω_i into $[0,1]^d$ so that the volume of $T_i(\Omega_i)$ is at least $1/d!$, then by scaling (2.5), and using the fact that

$$\sum_{i=1}^{k} |\Omega_i| = |\Omega \setminus (1-\delta/\delta_0)\Omega| = [1-(1-\delta/\delta_0)^d]|\Omega| \le d\delta/\delta_0,$$

we have for each $1 \le i \le k$,

$$N(\eta_i, \mathcal{C}_r(\Omega_i), \|\cdot\|_{L^p(\Omega_i)}) \le N(K\varepsilon, \mathcal{C}_r(T_i(\Omega_i)), \|\cdot\|_{L^p(T_i(\Omega_i))}),$$

where

$$K = 2^{-1/p}[2d^2(d!)^2\delta]^{\frac{1}{r}-\frac{1}{p}}.$$

Plugging into (2.13), we obtain

$$\begin{aligned} \log N(\varepsilon, \mathcal{C}_r(\Omega), \|\cdot\|_{L^p(\Omega)}) \le& (k+1)\log 4 + C_2 \delta^{-\frac{d}{2}-\frac{d^2}{r}} \varepsilon^{-d/2} \\ &+ \sum_{i=1}^{k} \log N(K\varepsilon, \mathcal{C}_r(T_i(\Omega_i)), \|\cdot\|_{L^p(T_i(\Omega_i))}). \end{aligned} \tag{2.14}$$

Now let $\mathcal{F}_k$ consist of all closed convex sets in $[0,1]^d$ with at most k faces and with volume at least $1/d!$, and define

$$g(k,\varepsilon) = \sup\left\{ \log N(\varepsilon, \mathcal{C}_r(\Omega), \|\cdot\|_{L^p(\Omega)}) \,\middle|\, \Omega \in \mathcal{F}_k \right\}.$$

For notational simplicity, we denote $M = C_2 \delta^{-\frac{d}{2}-\frac{d^2}{r}}$. Then (2.14) together with the fact that $(k+1)\log 4 \le 4k-4$ which follows from the fact that $k \ge 3$) implies

$$g(k,\varepsilon) + 4 \le M\varepsilon^{-d/2} + k[g(k+1, K\varepsilon) + 4]. \tag{2.15}$$

which is equivalent to

$$[g(k,\varepsilon)+4]\varepsilon^{d/2} \le M + \frac{k}{K^{d/2}}[g(k+1,K\varepsilon)+4](K\varepsilon)^{d/2}.$$

Now, we choose δ so that $K^{d/2} = 2k$. Then

$$M = C_2\delta^{-\frac{d}{2}-\frac{d^2}{r}} = C_3 k^{\frac{(r+2d)p}{r-p}}.$$

Thus,

$$[g(k,\varepsilon)+4]\varepsilon^{d/2} \le C_3 k^{\frac{(r+2d)p}{r-p}} + \frac{1}{2}[g(k+1,(2k)^{2/d}\varepsilon)+4]((2k)^{2/d}\varepsilon)^{d/2}.$$

Hence, for any positive integer m, we have

$$[g(k,\varepsilon)+4]\varepsilon^{d/2} \le C_3 \sum_{j=0}^{m-1} \frac{(k+j)^{\frac{(r+2d)p}{r-p}}}{2^j} + 2^{-m}[g(k+m,L_m\varepsilon)+4](L_m\varepsilon)^{d/2},$$

where

$$L_m = \prod_{j=0}^{m-1}(2k+2j)^{2/d}.$$

In particular, if we choose m to be the smallest integer so that $L_m\varepsilon \ge 1$, then $g(k+m, L_m\varepsilon) = 0$, and we obtain

$$g(k,\varepsilon) \le C_4 k^{\frac{(r+2d)p}{r-p}}\varepsilon^{-d/2}.$$

This finishes the proof of the upper bound with constant of the order k^γ with $\gamma = \frac{(r+2d)p}{r-p}$. □

2.6. **Upper Bound for Polytopes.** Note that if Ω is a convex polytope with v extreme points, then it has no more than $2v^{\lfloor d/2\rfloor}$ facets; see [10], Propositions 5.5.2 and 5.5.3, page 100. Therefore, we immediately obtain the upper bound with constant of the order $v^{\gamma\lfloor d/2\rfloor}$. We show that this estimate can be improved to $v^{\lceil d/2\rceil}$. Indeed, if Ω has v vertices, then it is known that Ω can be triangulated into $m = O(v^{\lceil d/2\rceil})$ many d-simplices; this is Corollary 2.3 of [13]; see also [2]. Thus, we can write $\Omega = \cup_{i=1}^m D_i$, where D_i are d-simplices. Because each D_i has only $(d+1)$-facets, by what we have proved above it follows that

$$\log N(\eta_i, \mathcal{C}_r(D_i), \|\cdot\|_{L^p(D_i)}) \le C_5|D_i|^{\frac{d}{2p}-\frac{d}{2r}}\eta_i^{-d/2},$$

where C_5 is a constant depending only on p, r, d. Now applying (2.12), with

$$\eta_i = 2^{-\frac{r-p}{p}}\left(\frac{|D_i|}{|\Omega|}\right)^{\frac{1}{p}-\frac{1}{r}}\varepsilon,$$

we immediately obtain

$$\begin{aligned}\log N(\varepsilon, \mathcal{C}_r(\Omega), \|\cdot\|_{L^p(\Omega)}) &\le \sum_{i=1}^m \log N(\eta_i, \mathcal{C}_r(D_i), \|\cdot\|_{L^p(D_i)}) \\ &\le C_6 m|\Omega|^{\frac{d}{2p}-\frac{d}{2r}}\varepsilon^{-d/2} \le C_7 v^{\lceil d/2\rceil}\varepsilon^{-d/2}.\end{aligned}$$

This proves Part (ii) of Theorem 1.1.

2.7. **General Upper Bound.** Fix $0<\varepsilon<1$; we choose smallest integer s so that $2^{-s}|\Omega|\le[2^{-1/p}\varepsilon]^{\frac{rp}{r-p}}|\Omega|$. By the definition of $S(t,\Omega)$, Ω contains $m_1\le S(1/2,\Omega)$ d-simplices $D_{1,i}$, $1\le i\le m_1$, so that the volume of $\Omega\setminus\cup_{i=1}^{m_1}D_{1,i}$ is at most $2^{-1}|\Omega|$, and the set $\Omega\setminus\cup_{i=1}^{m_1}D_{1,i}$ contains $m_2\le S(1/4,\Omega)$ d-simplices $D_{2,j}, 1\le j\le m_2$, so that the volume of

$$\Omega\setminus\cup_{i=1}^{2}\cup_{j=1}^{m_i}D_{i,j}$$

is at most $2^{-2}|\Omega|$. Continuing this way, we obtain a sequence of d-simplices $D_{i,j}, 1\le j\le m_i, 1\le i\le s$ that are packed in Ω so that the uncovered volume of Ω is at most $2^{-s}|\Omega|$. If we denote

$$\widehat{\Omega}_i=\cup_{k=1}^{i}\cup_{j=1}^{m_k}D_{k,j},$$

then for all $f\in\mathcal{C}_r(\Omega)$,

$$\int_{\Omega\setminus\widehat{\Omega}_s}|f|^p d\lambda\le|\Omega\setminus\widehat{\Omega}_s|^{1-\frac{p}{r}}\le\frac{\varepsilon^p}{2}|\Omega|^{1-\frac{p}{r}}.$$

Hence,

$$N(\varepsilon|\Omega|^{\frac1p-\frac1r},\mathcal{C}_r(\Omega),\|\cdot\|_{L^p(\Omega)})\le N(2^{-1/p}\varepsilon|\Omega|^{\frac1p-\frac1r},\mathcal{C}_r(\widehat{\Omega}_s),\|\cdot\|_{L^p(\widehat{\Omega}_s)}).\tag{2.16}$$

Next, we choose

$$\eta_{i,j}=2^{-1/p}\left(\frac{|D_{i,j}|}{\sum_{j=1}^{m_i}|D_{i,j}|}\cdot\frac{\alpha_i}{\sum_{k=1}^{s}\alpha_k}\right)^{\frac1p-\frac1r}\cdot\frac{\varepsilon}{2}|\Omega|^{\frac1p-\frac1r},$$

where

$$\alpha_i=(2^{-i}|\Omega|)^{1-\beta}[S(2^{-i},\Omega)]^{\beta},\quad\beta\equiv\frac{2pr}{2pr+(r-p)d}.$$

Using the fact that $\sum_{j=1}^{m_i}|D_{i,j}|\le 2^{-(i-1)}|\Omega|$, we have

$$\eta_{i,j}|D_{i,j}|^{\frac1r-\frac1p}\ge 2^{-1/p}\left(\frac{\alpha_i}{2^{-(i-1)}|\Omega|\sum_{k=1}^{s}\alpha_k}\right)^{\frac1p-\frac1r}\cdot\frac{\varepsilon}{2}|\Omega|^{\frac1p-\frac1r}.$$

Thus, together with the fact that $m_i\le S(2^{-i},\Omega)$, we have

$$\begin{aligned}&\sum_{j=1}^{m_i}\log N(\eta_{i,j},\mathcal{C}_r(D_{i,j}),\|\cdot\|_{L^p(D_{i,j})})\\ \le\;& S(2^{-i},\Omega)\cdot c\left[2^{-1/p}\left(\frac{\alpha_i}{2^{-i}|\Omega|\sum_{k=1}^{s}\alpha_k}\right)^{\frac1p-\frac1r}\cdot\frac{\varepsilon}{2}|\Omega|^{\frac1p-\frac1r}\right]^{-d/2}\\ =\;& c2^{\frac d2+\frac{d}{2p}}\left(\sum_{k=1}^{s}\alpha_k\right)^{\frac{(r-p)d}{2pr}}\alpha_i\cdot[\varepsilon|\Omega|^{\frac1p-\frac1r}]^{-d/2}.\end{aligned}$$

Therefore, by (2.12) and (2.16) we have,

$$\begin{aligned}&\log N(\varepsilon|\Omega|^{\frac1p-\frac1r},\mathcal{C}_r(\Omega),\|\cdot\|_{L^p(\Omega)})\\ \le\;&\log 4\sum_{i=1}^{s}S(2^{-i},\Omega)+c2^{\frac d2+\frac{d}{2p}}\left(\sum_{k=1}^{s}\alpha_k\right)^{1/\beta}\cdot[\varepsilon|\Omega|^{\frac1p-\frac1r}]^{-d/2}.\end{aligned}$$

Let $\gamma \equiv rp/(r-p)$. Note that $2^{-s} \le [2^{-1/p}\epsilon]^\gamma \le 2^{-(s-1)}$, and $S(t,\Omega) \ge S(2^{-i},\Omega)$ for $t \in [2^{-(i+1)}, 2^{-i}]$. Thus it follows that

$$\begin{aligned}
\sum_{i=1}^{s} \alpha_i &= \sum_{i=1}^{s} \left(2^{-i}|\Omega|\right)^{1-\beta} S(2^{-i},\Omega)^\beta \\
&= |\Omega|^{1-\beta} \sum_{i=1}^{s} 2^{-i} \left(\frac{S(2^{-i},\Omega)}{2^{-i}}\right)^\beta \\
&= 2|\Omega|^{1-\beta} \sum_{i=1}^{s} \int_{2^{-i-1}}^{2^{-i}} \left(\frac{S(2^{-i},\Omega)}{2^{-i}}\right)^\beta dt \\
&\le 2|\Omega|^{1-\beta} \sum_{i=1}^{s} \int_{2^{-i-1}}^{2^{-i}} \left(\frac{S(t,\Omega)}{t}\right)^\beta dt \\
&= 2|\Omega|^{1-\beta} \int_{2^{-s-1}}^{2^{-1}} \left(\frac{S(t,\Omega)}{t}\right)^\beta dt \\
&\le 2|\Omega|^{1-\beta} \int_{2^{-2}\cdot[2^{-1/p}\varepsilon]^\gamma}^{1} \left(\frac{S(t,\Omega)}{t}\right)^\beta dt.
\end{aligned}$$

Hence,

$$c2^{\frac{d}{2}+\frac{d}{2p}} \left(\sum_{k=1}^{s} \alpha_k\right)^{1/\beta} \cdot [\varepsilon|\Omega|^{\frac{1}{p}-\frac{1}{r}}]^{-d/2} \le c2^{\frac{d}{2}+\frac{d}{2p}+\frac{1}{\beta}} \left(\int_{\delta(\varepsilon)}^{1} \left(\frac{S(t,\Omega)}{t}\right)^\beta dt\right)^{1/\beta} \cdot \varepsilon^{-d/2},$$

where $\delta(\varepsilon) = 2^{-2} \cdot [2^{-1/p}\varepsilon]^\gamma$.

Similarly,

$$\sum_{i=1}^{s} S(2^{-i},\Omega) \le 2\int_{2^{-2}\cdot[2^{-1/p}\varepsilon]^\gamma}^{1} \frac{S(t,\Omega)}{t} dt.$$

Hence, we obtain

$$\begin{aligned}
&\log N(\varepsilon|\Omega|^{\frac{1}{p}-\frac{1}{r}}, \mathcal{C}_r(\Omega), \|\cdot\|_{L^p(\Omega)}) \\
&\le \quad C\int_{\delta(\varepsilon)}^{1} \frac{S(t,\Omega)}{t} dt + C\left(\int_{\delta(\varepsilon)}^{1} \left(\frac{S(t,\Omega)}{t}\right)^\beta dt\right)^{1/\beta} \cdot \varepsilon^{-d/2}
\end{aligned}$$

with $C = \max\left\{2\log 4, c2^{\frac{d}{2}+\frac{d}{2p}+\frac{1}{\beta}}\right\}$.

2.8. **Upper Bound for the Ball.** As a specific example, we consider the case when Ω is the closed unit ball in $\mathbb{R}^d$. We claim that there exists a simplicial approximation sequence $\{D_1, D_2, \ldots\}$ so that the corresponding simplicial approximation number $S(t,\Omega) = O(t^{-\frac{d-1}{2}})$.

Denote $I_d \equiv [-1,1]^d \supset \Omega$. For any integer k with $2^{2-k}\sqrt{d} < 1$, each facet of I_d can be divided into $2^{k(d-1)}$ closed $(d-1)$-cubes of side-length 2^{1-k}. Each of the these small $(d-1)$-cubes can be triangulated into no more than $d!$ closed $(d-1)$-simplices. Thus, the boundary of I_d can be triangulated into $m_k \le 2^{k(d-1)+1}dd!$ closed $(d-1)$-simplices, each of which has edge-length at most $2^{1-k}\sqrt{d}$. Let Δ_i, $1 \le i \le m_k$ be these simplices. For each $1 \le i \le m_k$, by connecting the origin with the d vertices of Δ_i, we obtain d line segments. These d line segments intersect the boundary of Ω at d points. Denote the convex hull of these d points by $\widetilde{\Delta}_i$.

Thus $\widetilde{\Delta}_i$ is a $(d-1)$-simplex with edge-length less than $2^{1-k}\sqrt{d}$. Let P_k be the convex hull of $\widetilde{\Delta}_i, 1 \le i \le m_k$. Then P_k is a simplicial sphere contained in Ω with $m_k \le 2^{k(d-1)+1}dd!$ facets. For each $x \in \Omega \setminus P_k$, we show that $\|x\| \ge \sqrt{1-2^{2-2k}d}$. Indeed, if we connect x and the origin, then the line segment from 0 to x intersects some facet $\widetilde{\Delta}_i$ of P_k, say at y. If $\|x\| < \sqrt{1-2^{2-2k}d}$, then $\|y\| < \sqrt{1-2^{2-2k}d}$. Let $C(y)$ be the spherical cap with a base disk centered at y. Then $C(y)$ has base radius larger than $2^{1-k}\sqrt{d}$. Because all the vertices of $\widetilde{\Delta}_i$ are within $2^{1-k}\sqrt{d}$ distance from y, these vertices all belong to the cap $C(y)$ above the base hyperplane. This is not possible because y belongs to the convex hull of these vertices. Therefore, for each $x \in \Omega \setminus P_k$, $\|x\| \ge \sqrt{1-2^{2-2k}d}$. Consequently, the volume of $\Omega \setminus P_k$ is at most $[1-(1-2^{2-2k}d)^{d/2}]|\Omega| \approx 2^{1-2k}d^2|\Omega|$.

Now, let k_1 be the smallest positive integer k such that $[1-(1-2^{2-2k}d)^{d/2}] \le 1/2$. Recall that the simplicial sphere P_{k_1} constructed above has $s_1 := m_{k_1}$ facets, each of which is a $(d-1)$-simplex. Let $D_1, D_2, \ldots, D_{s_1}$ be the convex hull of each facet with the origin. Then $D_1, D_2, \ldots, D_{s_1}$ is a triangulation of P_{s_1}. We define $S(t,\Omega) = s_1$ for $\frac{1}{2} \le t < 1$.

Suppose we have defined $S(t,\Omega)$ for $2^{-r} \le t < 1$, and the non-overlapping simplices $D_1, D_2, \ldots, D_{s_r}$ with the property

$$\cup_{i=1}^{s_r} D_i = P_{k_r}$$

where k_r is a positive integer k_r so that the volume of $\Omega \setminus P_{k_r}$ is at most $2^{-r}|\Omega|$. We choose k_{r+1} to be the smallest positive integer so that the volume of $\Omega \setminus P_{k_{r+1}}$ is at most $2^{-r-1}|\Omega|$. Thus, $k_{r+1} \ge k_r$. Note that the simplicial sphere $P_{k_{r+1}}$ constructed above has no more than $q := m_{k_{r+1}}$ facets, each of which is a $(d-1)$-simplex. Let $S_1, S_2 \cdots S_q$ be the convex hull of the origin with each of these facets. Then $S_1, S_2, \ldots, S_q$ is a triangulation of $P_{k_{r+1}}$. Note that each S_i intersects with no more than $d!$ facets of P_{k_r}. Indeed, each S_i is contained in the convex hull of the origin and a $(d-1)$ cube of side-length $2^{1-(k_{r+1})}$ that lies on a facet of $[-1,1]^d$, which is further contained in the convex hull of the origin and a $(d-1)$cube of side-length $2^{1-(k_r)}$ on a facet of $[-1,1]^d$. However, the latter convex hull intersects with no more than $d!$ facets of P_{k_r}. Thus, there exists a constant $c(d)$ depending only on d such that $S_i \setminus P_{k_r}$ can be triangulated into $c(d)$ d-simplices. Therefore, $P_{k_{r+1}} \setminus P_{k_r}$ can be triangulated into $c(d)q$ d-simplices. We define these simplices as $D_{s_r+1}, D_{s_r+2}, \cdots, D_{s_{r+1}}$, where $s_{r+1} = s_r + c(d)q$, and define $S(t,\Omega) = s_{r+1}$ for $2^{-r-1} \le t < 2^{-r}$

Next, we show that for this sequence of simplices, $S(t,\Omega) = O(t^{-\frac{d-1}{2}})$. Indeed, since for each $i = 1, 2, ..., k_i$ is the smallest integer k satisfying $[1-(1-2^{2-2k}d)^{d/2}] \le 2^{-i}$, we have $2^{-2k_i} \approx 2^{-i-1}/d^2$. Also note that $m_{k_i} \le 2^{k_i(d-1)+1}dd!$, which implies that $m_{k_i} \le k_d 2^{i(d-1)/2}$ for some constant k_d depending only on d. Consequently, for any $2^{-r-1} \le t \le 2^{-r}$, by the construction above

$$\begin{aligned} S(t,\Omega) &= m_{k_1} + c(d)m_{k_2} + \cdots + c(d)m_{k_{r+1}} \le c(d)\sum_{i=1}^{r+1} m_{k_i} \\ &\le c(d)k_d \sum_{i=1}^{r+1} 2^{i(d-1)/2} \le k'_d t^{-(d-1)/2}, \end{aligned}$$

for some constant k'_d depending only on d.

A direct computation of the integrals in Theorem 1.3 gives an upper bound of $\varepsilon^{-\frac{(d-1)p}{2}}$ when $p > \frac{d}{d-1}$; $\varepsilon^{-\frac{d}{2}}|\log\varepsilon|^{\frac{d+1}{2}}$ when $d = \frac{d}{d-1}$; and $\varepsilon^{-d/2}$ (from the second integral term) when $p < \frac{d}{d-1}$. This implies the estimate in the corollary.

2.9. **General Lower Bound.** By Lemma 2.1 and Lemma 2.2, we only need prove it for the case when Ω is contained in $[0,1]^d$ and has volume at least $1/d!$. Indeed, by Lemma 2.1, if $\Omega \subset \mathbb{R}^d$ is closed and convex, $\Omega \subset R$ for a box R with $|R| \le d!|\Omega|$. Let T be any affine transformation that maps R onto $[0,1]^d$, Then by Lemma 2.2, and the fact that $C_r(T(\Omega)) \supset C_\infty(T(\Omega))$, we have

$$\begin{aligned} N(|\Omega|^{\frac{1}{p}-\frac{1}{r}}\epsilon, \mathcal{C}_r(\Omega), \|\cdot\|_{L^p(\Omega)}) &= N((|\Omega|/|R|)^{\frac{1}{p}-\frac{1}{r}}\epsilon, \mathcal{C}_r(T(\Omega)), \|\cdot\|_{L^p(T(\Omega))}) \\ &\ge N(\epsilon, \mathcal{C}_\infty(T(\Omega)), \|\cdot\|_{L^p(T(\Omega))}). \end{aligned}$$

Thus it suffices to establish a lower bound for the case when Ω is contained in $[0,1]^d$ and has volume at least $1/d!$.

We choose a function f so that f is supported on $[0,1]^d$, with $0 \le f \le \frac{1}{20}$ and $\|f\|_1 \ge \frac{1}{80d}$. Furthermore, the Hessian matrix of f at every $(x_1, x_2, \ldots, x_d) \in [0,1]^d$ is a diagonal matrix with each entry bounded by 1. One such function is

$$f(x_1, x_2, \ldots, x_d) = \begin{cases} \frac{1}{20d}\sum_{i=1}^d \sin^3(\pi x_i) & \text{if } (x_1, x_2, \ldots, x_d) \in [0,1]^d \\ 0 & \text{if } (x_1, x_2, \ldots, x_d) \notin [0,1]^d \end{cases}.$$

For each fixed $0 < \varepsilon < (10d!)^{-2}$, and each $I = (i_1, i_2, \ldots, i_d) \in \mathbb{N}^d$, define

$$f_I(x_1, x_2, \ldots, x_d) = \varepsilon^2 \cdot f\left(\frac{x_1 - i_1\varepsilon}{\varepsilon}, \frac{x_2 - i_2\varepsilon}{\varepsilon}, \ldots, \frac{x_d - i_d\varepsilon}{\varepsilon}\right).$$

Then, f_I is supported on

$$B_I := [i_1\varepsilon, (i_1+1)\varepsilon] \times [i_2\varepsilon, (i_2+1)\varepsilon] \times \cdots \times [i_d\varepsilon, (i_d+1)\varepsilon]$$

with $0 \le f_I \le \frac{\varepsilon^2}{20}$, $\|f_I\|_1 \ge \frac{\varepsilon^2}{80d} \cdot \varepsilon^d$, and furthermore, the Hessian matrix of f_I at every $(x_1, x_2, \ldots, x_d) \in B_I$ is a diagonal matrix with each entry bounded by 1.

Denote

$$\mathcal{I} = \left\{ I \;\middle|\; I = (i_1, i_2, \ldots, i_d) \in \mathbb{N}^d, B_I \subset \Omega \right\}.$$

Let $\xi_I \in \{0,1\}$, $I \in \mathcal{I}$ be i.i.d. random variables with $\mathbb{P}(\xi_I = 1) = \mathbb{P}(\xi_I = 0) = 1/2$, and define the random function

$$F(x;\xi) = \sum_{I \in \mathcal{I}} \xi_I f_I(x).$$

Then for each realization of $\xi = (\xi_I)_{I\in\mathcal{I}}$, we have $0 \le F \le \frac{\varepsilon^2}{20}$, and the Hessian matrix of F is diagonal with each entry bounded by 1. Therefore, for each realization of ξ the function

$$G(x;\xi) = \frac{1}{d}\left(x_1^2 + x_2^2 + \cdots + x_d^2 - F(x;\xi)\right)$$

is convex and bounded by 1. Hence, $G(\cdot;\xi) \in \mathcal{C}_\infty([0,1]^d)$.

There are $2^{|\mathcal{I}|}$ realizations of $G(\cdot;\xi)$. Between two realizations, we define the Hamming distance

$$H(G(\cdot;\xi^{(1)}), G(\cdot;\xi^{(2)})) = \#\left\{ I \in \mathcal{I} \;\middle|\; \xi_I^{(1)} \ne \xi_I^{(2)} \right\}.$$

For $r = \lfloor |\mathcal{I}|/10 \rfloor$, consider the set

$$U(G(\cdot;\xi), r) = \left\{ G(\cdot;\xi^{(2)}) : \; H(G(\cdot;\xi), G(\cdot;\xi^{(2)})) \le r \right\}.$$

For each $G(\cdot;\xi)$, the set $U(G(\cdot;\xi), r)$ contains no more than

$$\sum_{k=0}^{r} \binom{|\mathcal{I}|}{k} \le 2^{9|\mathcal{I}|/10}$$

elements. Thus, by the pigeonhole principle, we can find $m \ge 2^{|\mathcal{I}|} \div 2^{9|\mathcal{I}|/10} = 2^{|\mathcal{I}|/10}$ realizations of $G(\cdot;\xi^{(k)})$, $1 \le k \le m$, such that for any $1 \le i < j \le m$, we have

$$H(G(\cdot;\xi^{(i)}), G(\cdot;\xi^{(j)})) \ge \lfloor |\mathcal{I}|/10 \rfloor.$$

Note that

$$\begin{aligned}
\int_\Omega \left| G(x;\xi^{(i)}) - G(x;\xi^{(j)}) \right| dx =& \frac{1}{d^2} \int_\Omega \sum_{I \in \mathcal{I}} |\xi_I^{(i)} - \xi_I^{(j)}| |f_I(x)| dx \\
\ge& \frac{1}{d} \sum_{I \in \mathcal{I}} |\xi_I^{(i)} - \xi_I^{(j)}| \frac{\varepsilon^2}{80d} \cdot \varepsilon^d \\
\ge& \frac{1}{d} \cdot \lfloor |\mathcal{I}|/10 \rfloor \cdot \frac{\varepsilon^2}{80d} \cdot \varepsilon^d. \qquad (2.17)
\end{aligned}$$

We show that the cardinality $|\mathcal{I}|$ of $\mathcal{I}$ is at least $\frac{1}{2d!}\varepsilon^{-d}$. Indeed, because $\Omega \subset [0,1]^d$ has volume at least $1/d!$, and Ω is convex, so the set $[0,1]^d \backslash \Omega_{\sqrt{d}\varepsilon}$ has volume at most $1 - 1/d! + 2d \cdot \sqrt{d}\varepsilon$. Thus, $[0,1]^d \backslash \Omega_{\sqrt{d}\varepsilon}$ contains no more than $\varepsilon^{-d} \cdot [1 - 1/d! + 2d \cdot \sqrt{d}\varepsilon]$ cubes B_I. Any cube $B_I \subset [0,1]^d$ that is not contained in $[0,1]^d \setminus \Omega_{\sqrt{d}\varepsilon}$ does not intersect with $[0,1]^\backslash \Omega$, thus must be contained in Ω. Since $[0,1]^d$ contains $\lfloor 1/\varepsilon \rfloor^d$ such cubes, and we conclude that Ω contains at least

$$\lfloor 1/\varepsilon \rfloor^d - \varepsilon^{-d} \cdot \left[1 - 1/d! + 2d \cdot \sqrt{d}\varepsilon \right] \ge \frac{1}{2d!} \varepsilon^{-d}$$

cubes provided that ε is small, say $\varepsilon < (10d!)^{-2}$.

Now plugging the inequality $|\mathcal{I}| \ge \frac{1}{2d!}\varepsilon^{-d}$ into (2.17), we obtain

$$\int_\Omega \left| G(x;\xi^{(i)}) - G(x;\xi^{(j)}) \right| dx \ge c\varepsilon^2,$$

for some constant c depending only on d. This implies that $\mathcal{C}_\infty([0,1]^d)$ contains

$$m \ge 2^{|\mathcal{I}|/10} \ge e^{c'\varepsilon^{-d}}$$

functions whose mutual $L^1(\Omega)$ distance is at least $c\varepsilon^2$. This implies that

$$\log N(\varepsilon, \mathcal{C}_\infty([0,1]^d), \|\cdot\|_{L^1(\Omega)}) \ge c'' \varepsilon^{-d/2}$$

for some $c'' > 0$ depending only on d.

Since $|\Omega| \ge \frac{1}{d!}$, for any $p \ge 1$, we have $\|\cdot\|_{L^p(\Omega)} \ge (d!)^{-\frac{p-1}{p}} \|\cdot\|_{L^2(\Omega)}$, this implies that

$$\log N(\varepsilon, \mathcal{C}_\infty([0,1]^d), \|\cdot\|_{L^p(\Omega)}) \ge c\varepsilon^{-d/2}$$

for some constant c depending on p and d, provided that $|\Omega| \ge \frac{1}{d!}$.

Together with the discussion at the beginning of this subsection, and the fact that bracketing entropy is bounded below by metric entropy we conclude that the lower bound statements of Theorem 1.1 are true.

2.10. **Lower Bound for the Ball.** The $(d-1)$-dimensional area of the unit sphere in $\mathbb{R}^d$ is $2\pi^{d/2}/\Gamma(d/2)$, while the $(d-1)$-dimensional area of a cap with height h is $(\pi^{d/2}/\Gamma(d/2))I_{2h-h^2}((d-1)/2, 1/2) \sim c_d h^{(d-1)/2}$ where $I_x(a,b)$ is the regularized incomplete beta function. Thus there exist $s \equiv \alpha_d h^{-(d-1)/2}$ disjoint spherical caps with height h. The $d-$dimensional volume of each spherical cap is $\beta_d h^{(d+1)/2}$. Let $x_1, \ldots, x_s$ be the spherical center of the caps. For each $1 \le i \le s$, we define a random function f_i on the closed unit ball such that

$$f_i(y) = \begin{cases} 0, & \langle y, x_i\rangle \le 1-h, \\ \xi_i \frac{\langle y, x_i\rangle - (1-h)}{h}, & \langle y, x_i\rangle > 1-h, \end{cases}$$

where ξ_i is either 0 or 1. Now f_i is convex on the closed unit ball, and supported on the $i-$th cap. Furthermore, since the caps are disjoint, the sum $f = \sum_{i=1}^s f_i$ is also convex and bounded by 1. There are 2^s different realizations of f. By the same argument as we used in the proof of the lower bound of Theorem 1.1, we can find a set W of $2^{s/2}$ functions in which any two functions f and g are different on at least $s/10$ caps.

On each cap where the two functions are defined differently, $|f-g| \ge 1/2$ the top half height of the cap which has a volume $\gamma_d h^{(d+1)/2}$. Consequently the L^p distance between any two functions $f, g \in W$ is at least

$$\frac{1}{2}(s/10 \cdot \gamma_d h^{(d+1)/2})^{1/p} \ge \delta_d h^{1/p}.$$

Letting $\delta_d h^{1/p} = \varepsilon$ we have

$$N(\varepsilon, \mathcal{C}_\infty(D), \|\cdot\|_{L^p(D)}) \ge \exp\left(C\varepsilon^{-(d-1)p/2}\right).$$

When $(d-1)p \le d$, the lower bound above should be replaced by the universal lower bound $\varepsilon^{-d/2}$ proved in the last section.

References

[1] E. M. Bronshtein, *$\epsilon-$entropy of convex sets and functions*, Siberian Mathematical Journal **17** (1976), 393–398.

[2] T. K. Dey and J. Pach, *Extremal problems for geometric hypergraphs*, Discrete and Computational Geometry **19** (1998), 473–484.

[3] D. Dryanov, *Kolmogorov entropy for classes of convex functions*, Constr. Approx. **30** (2009), no. 1, 137–153. MR 2519658 (2010h:41037)

[4] R. M. Dudley, *Uniform Central Limit Theorems*, Cambridge Studies in Advanced Mathematics, vol. 63, Cambridge University Press, Cambridge, 1999. MR 1720712 (2000k:60050)

[5] Fuchang Gao, *Entropy estimate for k-monotone functions via small ball probability of integrated Brownian motion*, Electron. Commun. Probab. **13** (2008), 121–130. MR 2386068 (2008m:60063)

[6] Fuchang Gao and Jon A. Wellner, *Entropy estimate for high-dimensional monotonic functions*, J. Multivariate Anal. **98** (2007), no. 9, 1751–1764. MR MR2392431

[7] Branko Grünbaum, *Convex polytopes*, second ed., Graduate Texts in Mathematics, vol. 221, Springer-Verlag, New York, 2003, Prepared and with a preface by Volker Kaibel, Victor Klee and Günter M. Ziegler. MR 1976856 (2004b:52001)

[8] A. Guntuboyina, *Covering numbers of l_p balls of convex functions and sets*, Constructive Approximation **xx** (2015), to appear, available at http://arxiv.org/pdf/1403.6922.pdf.

[9] Adityanand Guntuboyina and Bodhisattva Sen, *Covering numbers for convex functions*, IEEE Trans. Inform. Theory **59** (2013), no. 4, 1957–1965. MR 3043776

[10] Jiří Matoušek, *Lectures on discrete geometry*, Graduate Texts in Mathematics, vol. 212, Springer-Verlag, New York, 2002. MR 1899299 (2003f:52011)

[11] P. McMullen, *The minimum number of facets of a convex polytope*, J. London Math. Soc. (2) **3** (1971), 350–354. MR 0285990 (44 #3207)

[12] Robert Osserman, *Bonnesen-style isoperimetric inequalities*, Amer. Math. Monthly **86** (1979), no. 1, 1–29. MR 519520 (80h:52013)

[13] B. L. Rothschild and E. G. Straus, *On triangulations of the convex hull of n points*, Combinatorica **5** (1985), no. 2, 167–179. MR 815583 (87i:52021)

[14] Aad W. van der Vaart and Jon A. Wellner, *Weak Convergence and Empirical Processes*, Springer Series in Statistics, Springer-Verlag, New York, 1996, With applications to statistics. MR 1385671 (97g:60035)

Department of Mathematics, University of Idaho, Moscow, ID 83844-1103, USA
E-mail address: fuchang@uidaho.edu

Department of Statistics, University of Washington, Seattle, WA 98195-4322, USA
E-mail address: jaw@stat.washington.edu